Why Law, Economics, and Organization?

Oliver E. Williamson
University of California, Berkeley
owilliam@haas.berkeley.edu

This paper is a draft and should not be quoted without the permission of the author.

December 2000
Why Law, Economics, and Organization?

Oliver E. Williamson*

Occasional dissents notwithstanding, it is widely conceded that law and economics is a success story. There being general agreement that parsimony is a virtue in science, any effort to expand or otherwise reshape the law and economics enterprise needs to justify itself. My primary justification for the inclusion of organization is that reliance by law and economics on the orthodox theory of the firm-as-production function, which is a technological construction, has led to a truncated understanding of economic organization and has resulted in public policy error. Although there are reasons to believe that the worst such errors are behind us, the future will present new puzzles for which public policy error is a lurking concern. The second benefit that I associate with law, economics, and organization is educational: a more veridical theory of economic organization in which private ordering is featured will benefit the teaching of contract law and the training of what Ronald Gilson (1984) has referred to as “transaction cost engineers” in the law schools.

There are four propositions: economic organization is very complex; orthodox microeconomic theory, especially the theory of the firm-as-production function, fails to engage and/or misconstrues some of the main purposes served by economic organization; organization theory undergirds the theory of the firm-as-governance structure; and a wide range of contractual and organizational phenomena are better understood from the governance structure perspective. The critical concession, for law and economics scholars, is to come to terms with the fact that the orthodox theory of the firm was never designed with reference to and is often poorly suited to interpret the puzzles of firm and market organization.

I begin with a sketch of some of the earlier strains on antitrust law and economics. I then identify a series of key contributions from the organization theory literature which, when interpreted from a transaction cost economizing perspective, provide support for the theory of the firm-as-governance structure. The value added of this construction is then developed by
contrasting price theoretic with transaction cost theoretic ways of dealing with a variety of public policy issues. Additional applications to the subject of contract and legal education are then discussed. Concluding remarks follow.

1. Some Strains

As interpreted by Richard Posner, orthodox law and economics “has no, or at least few, aspirations to change economic theory” (1993, p. 82). Posner furthermore holds that “organization theory…[adds] nothing to economics that the literature on information costs has not added years earlier” (1993, p. 84). The first of these implies that the relation between law and economics is a one way street: the law is a supplicant that relies on whatever economics has to offer and has little or nothing to contribute in return. The second implies that economic orthodoxy can proceed heedless of the critiques and contributions of organization theory.

A growing number of economists take issue with the latter. The burgeoning study of the “economics of organization” turns on two propositions—organizations matter and organizations are susceptible to analysis. This work draws significantly upon the contributions of organization theory. As to the first, even supplicants might be expected to speak out when they observe that the emperor has no clothes.

The two pillars of textbook economic orthodoxy are the neoclassical theory of the firm and the neoclassical theory of consumer behavior. I am principally concerned with the first of these; behavioral law and economics takes exception mainly with the latter (Rabin, 1998, p. 11).

The neoclassical theory of the firm-as-production function is a technological construction in which the inner workings of firms (and other modes of organization) are scanted. Lionel Robbins set the agenda for most of the 20th century with his view (expressed in the 1930s) that “the economist does not interest himself in the internal arrangements within organizations but only what happens on the market” (Coase, 1992, p. 714). As Harold Demsetz has since put it, “It is a mistake to confuse the firm of economic theory with its real-world namesake. The chief mission of neoclassical economics is to understand how the price system coordinates the use
of resources, not to understand the inner workings of real firms” (1983, p. 377). Attention is thus focused on supply and demand, prices and output. All well and good as far as it goes. But what if economics is expected to interpret nonstandard and unfamiliar forms of contract and organization? What if economics is asked to give public policy advice on the puzzles and anomalies of firm and market organization?

Working out of the firm-as-production function setup did not afford a lot of leeway. The main explanation for “aberrant” practices and structures that could not easily be explained in technological terms (economies of scale and scope) became that of monopoly. Ronald Coase describes the prevailing state of confusion within industrial organization in the early 1970s as follows (1972, p. 67):

One important result of this preoccupation with the monopoly problem is that if an economist finds something—a business practice of one sort or another—that he does not understand, he looks for a monopoly explanation. And as in this field we are very ignorant, the number of ununderstandable practices tends to be rather large, and the reliance on a monopoly explanation frequent.

His advice for remedying this state of affairs was to develop a theory of the firm that is more attuned to the inner workings of firms, broadly in the spirit of his classic 1937 article on “The Nature of the Firm” (Coase, 1972, pp. 62-64). Inasmuch as organization theory takes the internal organization of firms very seriously, might there be lessons to be gleaned from this literature? Coase never commented on that possibility but did advise that “what is wanted in industrial organization is a direct approach to the problem. This would concentrate on what activities firms undertake, and would endeavor to discover the characteristics of the groupings of activities within firms…. In addition to studying what happens within firms, studies should also be made of the contractual relations between firms” (Coase, 1972, p. 73; emphasis added).

The potential benefits of a direct approach to the problem notwithstanding, launching such a project presumes that the theoretical base is really adequate. Of the myriad of activities
undertaken by firms, which are of first order importance and why? Willard Hurst’s densely descriptive history of the lumber industry in Wisconsin (Hurst, 1964) is illustrative. Because it lacks a theoretical framework, the priorities are obscure and the text, for some, is nearly “unreadable” (Posner, 1993, p. 74). If the theory of the firm as described by Coase (1937) was tautological (Alchian and Demsetz, 1972) and not really comparative (see Section 3), what to do?

2. Organization Theory: Proximate and Ultimate Lessons

Although important new developments in organization theory were beginning to surface in the 1930s (Merton, 1936; Barnard, 1938), it took another decade for the lessons of Barnard to be digested (Simon, 1947) and yet another decade before the main contributions to the new field of organization theory were integrated (March and Simon, 1958). More, however, was needed. Organization theory and economics needed to be joined and the tensions between them worked out. The interdisciplinary program of social science training and research at Carnegie Tech in the 1950s and 1960s made considerable headway with these purposes.²

The behavioral theory of the firm (Cyert and March, 1963) stands out as the apogee of these Carnegie efforts. This work was much more influential in the field of organization theory than in economics, however. That is because it dealt with more microanalytic phenomena than were of interest to most economists (e.g., predicting department store prices to the exact penny) and because it worked out of a myopic setup in which local search, trial-and-error learning, and crisis management were featured.

Albeit influential, the behavioral theory of the firm does not exhaust the actual and potential lessons of organization theory for economics. Many of the pertinent lessons are discerned, however, only upon going beyond the proximate lessons to work out the ultimate lessons. This entails taking the logic of economic organization “to completion”—which is congruent, I think, with the “rational spirit” of the economist to which Kenneth Arrow (1974, p. 16) refers and with John Maynard Keynes’ description of economics as “a method rather than a
doctrine, a technique of thinking, which helps its possessor to draw correct conclusions” (quoted in Feiwel, 1987, p. v). George Schultz elaborates as follows: “my training in economics has had a major influence on the way I think about public policy tasks, even when they have no particular relationship to economics. Our discipline makes one think ahead, ask about indirect consequences, take note of variables that may not be directly under consideration” (1995, p. 1).

Because organization theory was unfamiliar terrain to most economists and because few economists perceived value-added in the proximate lessons of organization theory, a yawning gap separated these two. Three moves were needed for the lessons of organization theory to be incorporated within economics: (1) an identification of the key contributions, (2) choice of a productive lens, and (3) an effort to push the logic to completion. I use the lens of transaction cost economizing to interpret the key contributions from organization theory and to take the logic to completion.

On my reading, organization theory helps to reshape the research agenda by taking issue with orthodoxy in the following significant respects (Williamson, 1993a, 2000): (1) the description of human actors, (2) the central problem of economic organization, (3) process transformations, (4) unit of analysis, (5) discrete structural analysis, and (6) embeddedness.

2.1 human actors

If “Nothing is more fundamental in setting our research agenda and informing our research methods than our view of the nature of the human beings whose behavior we are studying” (Simon 1985, p. 303), then social scientists should be prepared to name the key attributes of human actors. Both the condition of cognition and self-interestedness need to be addressed.

Simon early took the position that the hyperrationality assumption out of which orthodox economics works should be supplanted by the less demanding cognitive condition of bounded rationality—according to which human actors are intendedly rational but only limitedly so. The issue of self-interest has received less attention, but Simon has described it as “frailty of motive”
and has also featured docility, identification, and loyalty in describing human behavior in organization (Simon 1991, 1997).

(a) proximate lessons

The main lesson for Simon of supplanting hyperrationality with bounded rationality is to replace “the goal of maximizing with the goal of satisficing, of finding a course of action that is ‘good enough’” (Simon, 1957a, pp. 204-205), while the chief lesson of ascribing “frailty of motive” to human actors (Simon, 1985, p. 303) is to describe behavior as being mainly benign. Most people will do what they are asked to do and some will do more (Simon, 1997, p. 35). In that event, organizational behavior can be predicted from knowledge of the operative routines (Cyert and March, 1963). Strategic behavior, in this scheme of things, is given short shrift.

(b) ultimate lessons

Transaction cost economics agrees that our view of the human beings whose behavior we are studying has profound ramifications for the research agenda. It also concurs that human actors are subject to bounded rationality. Rather, however, than dwell on the analytical apparatus to be employed (that is, maximizing, satisficing, game theory, or whatnot), TCE asks instead what key conceptual moves in orthodoxy are disallowed upon admitting to limited cognitive competence. The TCE response to this query is this: the chief lesson of bounded rationality for the study of economic organization is that all complex contracts are unavoidably incomplete.

The combination of incomplete contracts (by reason of bounded rationality) with unreliable reporting (by reason of opportunism, see below) undermine the idea that common knowledge between the two parties to a contract suffices to annihilate ex post contracting problems. Because it is implausible to ascribe common knowledge to ultimate arbiters of disputes (the courts), common knowledge between the principals does not preclude costly maladaptation and ex post bargaining (Williamson, 1975, pp. 31-37). The upshot is that, contrary to the usual game theoretic setups, costless bargaining cannot be invoked to effect
efficiency in the ex post contract implementation interval. Instead, the adaptive efficacy of alternative modes of organization (market, hybrid, hierarchy, etc.) in ex post governance respects now need to be discovered and factored into the comparative institutional calculus.

My reference to opportunism stands in contrast to Simon’s description of self-interest as frailty of motive. This does not dispute that most people will do what they say (and some will do more) without self-consciously asking whether the effort is justified by expected discounted net gains. If they slip, it is a normal friction and often a matter of bemusement. The proposition that routines describe the behavior of most individuals in organizations most of the time plainly contemplates such (nonstrategic) benign behavior.

But while accurate descriptions of what is going on “most of the time” are important, much of what is interesting about human behavior in general and organizations in particular has reference not to routines but to exceptions. Faced with unanticipated disturbances for which an incomplete contract makes inadequate or incorrect provision (by reason of gaps, errors, and omissions), such disturbances will push the parties to an incomplete contract off of the contract curve. Strategic considerations now come into play if, rather than frailty of motive, opportunism is the operative condition. The aforementioned maladaptation therefore results.

In addition to bounded rationality and opportunism, TCE also describes human actors with reference to foresight. Rather than myopia or omniscience, TCE postulates “feasible foresight.” The earlier quotes from Keynes and Schultz are pertinent. So also is Richard Dawkins’ observation that the “capacity to simulate the future in imagination…[saves] us from the worst consequences of the blind replicators” (1976, p. 200). Practitioners, consultants and public policy analysts who are alert to the lessons of experience and who possess the skills for and practice the art of feasible foresight will look ahead, identify potential contractual hazards, uncover the mechanisms through which the hazards operate, work out the ramifications for governance, and fold these back into the ex ante design of the contract—by introducing
contractual safeguards or, as appropriate, substituting an alternative mode of governance (as, for example, moving from a market mediated transaction to hierarchy).  

2.2 central problem: adaptation

Adaptation is taken to be the central problem of economic organization. But whereas autonomous adaptation in the market in response to changes in relative prices is what economics had long emphasized, organization theory featured cooperative adaptation of a “conscious, deliberate, purposeful” kind (Barnard, 1938, p. 4), working through administration.

(a) proximate lesson

The proximate lesson for economics is to be more respectful of the benefits of purposeful planning and coordination that accrue to internal organization.

(b) ultimate lesson

Interestingly, both the economist Friedrich Hayek (1945) and the organization theorist Chester Barnard (1938) are in agreement that adaptation is the central problem of economic organization. The adaptations to which they have reference, however, differ. Hayek had reference to the adaptations of autonomous economic actors who adjust spontaneously to changes in the market (mainly as signaled by changes in relative prices). By contrast, Barnard appealed to intentionality. He featured cooperative adaptations made by economic actors with the assistance of hierarchy within firms.

As against the orthodox propensity to take firm and market organization as given (determined largely by technology), TCE holds that choice of mode should be derived by recognizing that the adaptive needs of transactions (in autonomous and cooperative respects) vary with the attributes of transactions and that the adaptive capacities of alternative modes of governance also differ. The upshot is that efficiency gains are realized by aligning transactions with governance structures so as to effect an economizing outcome. Pushing the logic of autonomous and cooperative adaptation to completion thereby leads to a predictive theory of economic organization (Williamson, 1991).
2.3 process matters

If alternative modes of organization differ because of intrinsic process differences among them, and if organization, like the law, has a life of its own, then both key process differences and intertemporal process transformations need to be uncovered and the ramifications for the study of comparative economic organization wrung out. Process differences are discussed in 2.5, intertemporal process transformations—of which unintended consequences and the imperatives of bureaucracy are among the more important (March and Simon, 1958)—are discussed here. Students and practitioners who lack awareness of the chief intertemporal propensities of organization will necessarily fail to make allowance for real but unrecognized costs and benefits.

(a) proximate lessons

The knee-jerk response, if something in the firm goes wrong, is to fix it by introducing new rules and regulations—what March and Simon refer to as “demands for control” (1958, pp. 35-47). Such a “machine model” of organization is too simple, however, if human actors do not respond in a docile way and it is impossible to write a complete “repertoire of response programs…[for which] there is a unique stimulus or cure for each such program” (March and Simon, 1958, p. 34). Intended effects will be incompletely realized if they are attended by unanticipated consequences. In that event, the naïve machine model must give way to a more complex model in which human beings are strategic actors.

(b) ultimate lessons

TCE concurs that unanticipated consequences are important but takes exception with the view that firms are a mess (Gibbons, 2000) because they deviate from the neoclassical ideal. It being the case that all feasible forms of organization are flawed, what useful purpose is served by describing any or all as a mess? Markets, hybrids, firms, bureaus, etc. are simply alternative modes of governance with distinctive strengths and weaknesses. The need is to uncover the strengths and weaknesses of each. Organization theory is pertinent to all.
That internal organization has a life of its own has been evident to sociologists of organization for a long time, including the idea that taking the logic of economic organization to completion requires that the ex ante organizational design ramifications of these regularities be worked out. Robert Michels’ 1911 book on Political Parties focused on the intertemporal transformations that regularly attended democratic efforts at political organization. The most important such intertemporal transformation is summarized by the famous Iron Law of Oligarchy: “It is [hierarchical] organization which gives birth to the dominion of the elected over the electors, of the mandatories over the mandators, of the delegates over the delegators. Who says organization, say oligarchy” (Michels, 1962, p. 365). Michels traced the source of these oligarchical tendencies to “the nature of the human individual,…the nature of the political struggle,…and the nature of organization” (1962, p. 6).

Michels, moreover, had a very farsighted view of his findings. “The sociologist should aim…at the dispassionate exposition of tendencies and counter-operating forces, of reasons and opposing reasons, at the display, in a word, of the warp and the woof of social life” (Michels, 1962, p. 6). Unless we are alert to the intertemporal propensities of organization, we will be needlessly victimized by them: “nothing but a serene and frank examination of the oligarchical dangers of democracy will enable us to minimize these dangers” (1962, p. 370). Thus although the oligarchical propensities of democratic organization may have been poorly understood by academics and some practitioners until Michels clarified the issue, the lurking hazards of oligarchy should no longer come as a surprise. Today’s organizational designers presumably take the Iron Law of Oligarchy into account in the initial design calculus.

Philip Selznick characterized “Michels’ theory about democratic organization…as a special case of the general recalcitrance of the human tools of action. The tendency for goals to be subverted through the creation of new centers of interest and motivation inheres in all organizations” (1950, p. 162; emphasis added). The study of “unanticipated consequences” of all kinds—of which oligarchy is but one example—thus describes the larger research agenda.
TCE responds in a three-part way. First, be alert to all of the significant unanticipated consequences and bureaucratic propensities that students of internal organization uncover. Second, push beyond to discover the mechanisms through which these operate, whereupon the challenge is to work out the ex ante design ramifications such that unwanted effects can be mitigated (also wanted but initially unrecognized benefits can be enhanced) in cost-effective degree.

The added burdens of bureaucracy that accrue upon making rather than buying a good or service are pertinent in this respect. As discussed in 2.5, some of these “burdens” are immediate and are the foreseeable consequence of moving from high to lower-powered incentives. Others, however, are the result of more subtle internal political processes and show up with delay (Williamson, 1985, Chap. 6). Because it can be very difficult to reverse internal procurement decisions, it is important to recognize the tradeoffs between immediate benefits and delayed costs in the initial decision calculus.

Intertemporal tradeoffs between immediate benefits and future burdens also show up in the Fundamental Transformation, whereby a large numbers bidding competition is (sometimes) transformed into a small numbers supply relation during contract execution and at the contract renewal interval (Williamson, 1985, pp. 61-63). This is discovered by pushing the logic of incomplete contracting to completion. As discussed below, the Fundamental Transformation has pervasive ramifications for comparative economic organization.

2.4 unit of analysis

Moving from the firm as black box to an examination of the inner workings of firms entails choice and dimensionalization of a more microanalytic unit of analysis. One candidate is the concept of role, but Simon notes that this unit of analysis “has never been given sufficiently precise definition” (1957b, p. xxx).
The proximate lesson for economics is to choose and operationalize a unit of analysis that engages the decision making process. Simon recommends that the decision premise be made the unit of analysis and avers that “Behavior can be predicted…when the premises of the decision are known (or can be predicted) in sufficient detail” (Simon 1957b, p. xxx). A related but more composite variant of the decision premise is to treat the “routine”—the decision rules and procedures—as the unit of analysis (Cyert and March, 1963).

(b) ultimate lesson

Although the decision premise has been used in cognitive science (Newell and Simon, 1972) as the unit of analysis, it has never been shown to have general purpose application as an operational unit of analysis in organization theory. The routine as the unit of analysis has fared better within organization theory (Cyert and March, 1963) and has been adopted by both evolutionary economics (Nelson and Winter, 1982) and the core competence perspective (Teece and Pisano, 1994). The routine has never been operationalized, however, by either of these latter literatures.

TCE adopts the purposive perspective of John R. Commons in naming a unit of analysis: “The ultimate unit of activity…must contain in itself the three principles of conflict, mutuality, and order. This unit is a transaction” (Commons, 1932, p. 4). Not only does transaction cost economics concur that the transaction is the basic unit of analysis, but it views governance as the means by which to infuse order, thereby to mitigate conflict and realize mutual gains.

Naming a unit of analysis is always easier than providing operational content. Of the countless ways by which transactions differ, which are consequential? Where does the action reside? These questions were never asked, hence never answered, by older style institutional economics.

Transaction cost economics delimits the dimensions for describing transactions by focusing on factors that are responsible for the breakdown of the ideal transaction in law and economics—according to which contracts take place between faceless economic actors,
whence continuity is unimportant and the identity of the parties does not matter. Relevant attributes for describing transactions between parties where identity does matter include asset specificity in its various forms (which gives rise to bilateral dependency), uncertainty (for which coordinated adaptations to disturbances may be needed), and frequency (which has a bearing on the value of preserving a continuing relation and on the incentive to incur the cost of specialized governance).

2.5 discrete structural analysis

The term discrete structural analysis was introduced into the study of comparative economic organization by Simon, who observed that (1978, pp. 6-7):

As economics expands beyond its central core of price theory, and its central concern with quantities of commodities and money, we observe in it…[a] shift from a highly quantitative analysis, in which equilibration at the margin plays a central role, to a much more qualitative institutional analysis, in which discrete structural alternatives are compared….

Such analyses can often be carried out without elaborate mathematical apparatus or marginal calculation. In general, much cruder and simpler arguments will suffice to demonstrate an inequality between two quantities than are required to show the conditions under which these quantities are equated at the margin.
(a) proximate lessons

One proximate lesson is that discrete structural analysis is simple to implement. A second is that moving from one mode of organization to another is attended by discontinuities.

(b) ultimate lessons

Because the mathematical apparatus of marginal analysis is actually easy to implement, economists can be thought of as analytical satisficers: they use workable apparatus that (often) is “good enough.” The real import of discrete structural analysis is different. It resides in the proposition that the movement from one organizational form to another experiences discontinuities. The question thus turns to an examination of the factors that are responsible for these discontinuities.

The “employment relation” is one answer (Coase, 1937; Simon, 1951; Masten, 1988), but that mainly serves to distinguish classical markets from firms. Where do hybrid modes of contracting and where do public bureaus fit in? Employment also focuses on merely one feature of organization whereas recent work on the economics of organization (Williamson, 1988, 1991; Milgrom, Qian, and Roberts, 1991; Holmstrom and Milgrom, 1993) holds that discreteness is explained by the fact that alternative modes of organization display a complementary cluster of features.

Specifically, TCE holds that each generic mode of governance is defined by a syndrome of internally consistent attributes—of which incentive intensity, administrative controls, and the contract law regime are especially important. With reference to contract law, the argument is that each generic mode of governance is supported by a different contract law regime. In particular, the implicit law of internal organization is that of forbearance. Thus whereas courts routinely grant standing to interfirm disputes over prices, the damages to be ascribed to delays, failures of quality, and the like, courts will refuse to hear disputes between one internal division and another over identical technical issues. Access to the courts being denied, hierarchy is its
own court of ultimate appeal, whereupon firms have access to fiat that interfirm contracting does not.

More generally, the alignment differences by which firm and market are distinguished are these:

1. incentive intensity: the high-powered incentives of markets give way to low-powered incentives in firms;

2. administrative controls: as compared with markets, firms are supported by a more extensive array of administrative rules and procedures, to include accounting and auditing, as well as the supports of informal organization;

3. contract law: the contract law of markets is legalistic and relies on court ordering whereas, as described above, the contract law of internal organization is that of forbearance; and

4. adaptation: by reason of these differences, markets enjoy the advantage in effecting autonomous adaptations whereas the advantage accrues to firms in effecting cooperative adaptations.

As hitherto discussed, refutable implications accrue upon aligning transactions with governance structures so as to effect an economizing result.

2.6 informal organization

Barnard argued that formal and informal organization always and everywhere coexist (1938, p. 20) and that informal organization contributes to the viability of formal organization in three significant respects: “One of the indispensable functions of informal organization in formal organizations…[is] that of communication…. Another function is that of maintaining the cohesiveness in formal organizations through regulating the willingness to serve and the stability of objective authority. A third function is the maintenance of the feeling of personal integrity, of self-respect, and independent choice” (Barnard 1938, p. 122).
(a) proximate lessons

The informal effects to which Barnard refers occur spontaneously. Network forms of organization (Podolny and Page, 1998), sometimes with the assistance of informal ethnic ties (Saxenian, 2000), are especially important in this respect.

(b) ultimate lessons

Because “ongoing social relations” (Granovetter, 1985) within firms and between firms differ, these differences should presumably be taken into account in the decision to use market or hierarchy. Also, the efficacy of informal organization varies with the choice of internal organizational form (Chandler, 1966; Kreps, 1990) and is not equally valued for all transactions. These factors need to be worked out more fully and factored into the governance calculus. Interdependencies among transactions of an “economics of atmosphere” kind also arise (Williamson, 1975, 1993b).

2.7 other

As a scan of the table of contents of any leading text (such as W. Richard Scott’s text on Organizations: Rational, Natural, and Open Systems (1998)) reveals, organization theory is a huge subject and deals with many issues not addressed above. My suggestion for incorporating added insights from this literature is to follow the plan set out herein: describe the hitherto neglected issue and its proximate effects, then use the lens of transaction cost economizing to work out the ramifications for comparative economic organization.

3. The Theory of the Firm-as-Governance Structure

Theories of the firm of neoclassical, behavioral, agency theory, evolutionary, and transaction cost kinds can be distinguished. Of these five, the behavioral theory of the firm (Cyert and March, 1963) hews most closely in spirit to the proximate lessons set out above. Neoclassical theory makes provision for none⁶ and agency theory provides only for asymmetric information and opportunism. Transaction cost economics makes provision for all, but the main
lessons to be gleaned upon bringing economizing lens to bear are what I refer to as the ultimate lessons. The upshot is that there is a reciprocal relation between TCE and organization theory: each should inform the other.

Two of the crucial moves in the transaction cost treatment of economic organization are that economizing on transaction costs is an important and pervasive concept and that the costs and benefits to be ascribed to alternative modes of organization need to be worked up comparatively.

3.1 doing comparative institutional analysis

It is much easier to identify a cost than to establish a comparative cost difference between one mode of governance and another. For example, firms incur bureaucratic costs. This does not, however, imply that bureaucratic costs vanish upon mediating a transaction by the market rather than internally. It merely means that two firms rather than one firm are engaged in the exchange. Thus the political and managerial mechanisms to which added bureaucratic costs accrue upon taking a transaction out of the market and organizing it internally need to be identified and the ramifications worked out rather than asserted (Williamson, 1985, pp. 135-153). This is an exercise in comparative microanalytics.

Or consider the main argument made by Coase for taking transactions out of markets and organizing them within firms. According to Coase, “The main reason why it is profitable to establish a firm would seem to be that there is a cost of using the price mechanism, the most obvious…[being] that of discovering what the relevant prices are” (Coase, 1937, p. 391). That sounds plausible, but is it truly comparative? How is it that internal procurement by the firm avoids the costs of price discovery?

The “obvious” answer is that sole-source internal supply avoids the need to consult the market about prices because internal accounting prices of a formulaic kind (for example, of a cost-plus kind) can be used to transfer a good or service from one internal stage to another. If, however, that is the source of the advantage, the obvious lesson is to apply this same practice to
outside procurement. The firm simply advises its purchasing office to turn a blind eye (or a deaf ear) to the market by placing orders, period by period, with a qualified sole-source external supplier who agrees to sell on cost-plus terms. In that event, firm and market are put on a parity in price discovery respects—which is to say that the price discovery burden that Coase ascribes to the market does not survive comparative institutional scrutiny.

Coase and others might reply that the hazards of cost-plus procurement within and between firms are different. I do not disagree (indeed, I have addressed the comparative efficacy of cost-plus procurement within and between firms elsewhere (Williamson, 1985, pp. 153-155)). Note, however, that the real explanation now resides in identifying and explicating underlying governance structure differences between firm and market. Furthermore, these governance structure differences need to be linked to the attributes of transactions, since markets are displaced selectively rather than uniformly. A predictive theory of firm and market organization needs to uncover these differences and work out the logic of efficient alignment.

3.2 the organization theory bridge

Such criticisms of Coase do not in the least disparage the magnitude of his classic 1937 article. Coase was simply ahead of his times. Many of the requisite contributions from organization theory were not yet in place.

The theory of the firm-as-governance structure that is herein described is responsive to follow-on developments in organization theory in the following respects: It (1) works out of an incomplete contracting setup (bounded rationality) in which (2) hazards arise between bilaterally dependent parties (asset specificity) when they are (3) pushed off of the contract curve by disturbances (uncertainty), whereupon (4) the parties resort to haggling (opportunism). Such transactions invite (5) farsighted parties to relieve conflict and (6) promote cooperative adaptation by embedding the transactions in either credible contracting or hierarchical governance
structures, where (7) these two governance alternatives differ in discrete structural ways, of which (8) the comparative burdens of bureaucracy and differential access to fiat are two.

To those who might (understandably) complain that those are too many moves, my responses are: Which moves can be excised? Does TCE yield predictions that are corroborated by the data?

3.3 discriminating alignment: a recapitulation

The aforementioned discriminating alignment hypothesis—according to which transactions, which differ in their attributes, are aligned with governance structures, which differ in their cost and competence, in a transaction cost economizing way—is the principal TCE engine for deriving refutable implications. Note in this connection that TCE subscribes to Nicholas Georgescu-Roegen’s view that the “purpose of science in general is not prediction, but knowledge for its own sake,” yet prediction is “the touchstone of scientific knowledge” (1971, p. 37). Understanding is the object. Because, however, there are many plausible theories, there is a need to sort the sheep from the goats. Sooner or later, candidate theories of organization must go beyond ex post rationalization by advancing predictions and confronting the data.\(^8\) Note for this purpose that TCE invites contact with the data by naming the attributes of transactions (with special emphasis on asset specificity, uncertainty, and frequency) and the syndromes of attributes that describe alternative modes of governance (incentive intensity, administrative controls, contract law regime, and adaptation).

3.4 the firm-as-governance structure: a heuristic display\(^9\)

As described above, the key purpose of organization on which transaction cost economics focuses is that of adaptation, of which two kinds were distinguished: autonomous adaptation in response to changes in relative prices, and cooperative adaptation accomplished through administration. Not only do transactions have differing “needs” for adaptations of these
two kinds, but governance structures have differing capacities to supply autonomous and cooperative adaptation.

The simple contractual schema shown in Figure 1 is a heuristic illustration. Thus assume that a firm can make or buy a component and assume further that the component can be supplied by either of two technologies. One is a general purpose technology, the other a special purpose technology. The special purpose technology requires greater investment in transaction-specific durable assets and is more efficient for servicing steady-state demands. Steady-state, however, is an analytical convenience: most contracts are implemented under conditions of uncertainty for which adaptation to disturbances is needed. Because an incomplete contract is often silent or has otherwise made inadequate provision for some of these adaptations, contractual conflicts between bilaterally dependent parties (that is, those for which continuity has value) prospectively arise. Thus although mutual gains will be realized upon costlessly restoring a position on the contract curve, each party may posture and make opportunistic representations over the division of gains. Costly delays and imperfect adaptations result.

Using $h$ as a measure of contractual hazards, transactions that use the general purpose technology are ones for which $h = 0$. Autonomous adaptation in a competitive market suffices because the parties are faceless. If instead transactions use the special purpose technology, an $h > 0$ condition exists. Assets here are specialized, whence productive values would be sacrificed if $h > 0$ transactions were to be prematurely terminated. Such bilaterally dependent parties have incentives to promote continuity and safeguard investments. Coordination (cooperative adaptation) thus comes to the fore.

Let $s$ denote the magnitude of any such safeguards. An $s = 0$ condition is one in which no safeguards are provided; a decision to provide safeguards is reflected by an $s > 0$ result. Safeguards can take either of two forms. One would be to provide interfirm contracts with added support: penalties to deter breach are introduced, added information disclosure is provided, and
specialized dispute settlement machinery (e.g., arbitration) is devised. This is the credible interfirm commitment option. A second would be to take transactions out of markets and organize them under unified ownership where hierarchy (to include fiat) is used to effect coordination.

Node A corresponds to the ideal transaction in law and economics: there being an absence of dependency \((h = 0)\), prices are set competitively in the market (by supply and demand) and, in the event of contractual breakdown, the courts award damages. Node B poses unrelieved contractual hazards, in that specialized investments are exposed \((h > 0)\) for which no safeguards \((s = 0)\) have been provided. Such hazards will be recognized by farsighted players, who will price out the implied risks. Nodes C and D are those for which additional contractual support has been provided \((s > 0)\), either in the form of contractual safeguards (node C) or unified ownership (node D).

In the event that costly breakdowns continue in the face of best bilateral efforts to craft safeguards at node C, the transaction may be taken out of the market and organized under unified ownership (vertical integration) instead. Inasmuch, however, as added bureaucratic costs accrue upon taking a transaction out of the market and organizing it internally, internal organization is usefully thought of as the organization form of last resort: try markets, try hybrids, and have recourse to the firm only when all else fails. Node D, the firm, thus comes in only as higher degrees of asset specificity and added uncertainty pose greater needs for cooperative adaptation.

4. Applications to Public Policy

Node A excepted, which is the ideal transaction in law and economics to which I referred previously, the neoclassical and transaction cost approaches to firm and market organization plainly differ. These differences are due to the broader conception of economic organization out of which transaction cost economics works (where alternative modes of organization are described as governance structures, to which the lessons of organization theory apply), which
differences have ramifications for public policy toward business. Neoclassical and transaction cost interpretations of nonstandard and unfamiliar contracting practices and organizational structures are compared and contrasted here. The overarching difference is this: orthodox economics is more imperial in that it imposes a price theoretic interpretation on the phenomena in question, whereas transaction cost economics is more curious and asks the question “What’s going on here?” The TCE action is in the details of transactions on the one hand and governance structures on the other, which is closer in spirit to organization theory.

4.1 vertical integration/vertical market restraints

Orthodox explanations for integration (backward, forward, or lateral) of the firm-as-production function kind invoke considerations of technology, inefficient factor proportions that result from double-marginalization (McKenzie, 1951), and/or distortions that arise from government imposed quotas or sales taxes.

Joe Bain’s treatment of thermal economies, recently repeated by Daniel Spulber (1999, p. 270), is illustrative of technological reasoning (1968, p. 381):

…the cases of clear economies of integration generally involve a physical or technical aspect of the processes in a single plant. A classic case is that of integrating iron-making and steel-making to effect a saving in full costs by eliminating a reheating of iron before it is fed to a steel furnace. Where integration does not have this physical or technical aspect—as it does not, for example, in integrating the production of assorted components with the assembly of those components—the case for cost savings from integration is much less clear.

As a technological matter, however, the thermal economies to which Bain and Spulber refer actually require only that the two stages be located in close proximity to one another. That the two stages be placed under unified ownership is not implied. If, therefore, economies somehow accrue to the unified ownership of these two stages (that is, the relation between the
two stages is better mediated by hierarchy rather than by market), this must be due to other—possibly transactional—rather than technological reasons.

Transaction cost economics thus looks behind apparent explanations (such as price discovery or thermal economies) to see if they withstand comparative institutional scrutiny. It also asks whether outside procurement poses interfirm contractual hazards for which cost-effective relief will be realized upon taking the transaction in question into the firm (added bureaucratic costs notwithstanding). Specifically, the progressive buildup of contractual complications, as discussed in conjunction with the simple contractual schema in Figure 1, is what mainly explains successive moves from ideal market to hybrid to hierarchy.

So what about vertical market restrictions? How are these to be understood? For starters, vertical market restrictions can be interpreted as a decision to remain at node C rather than move to node D. The transaction in question is one to which hazards accrue \((h > 0)\) for which cost-effective safeguards are needed \((s > 0)\). If most of the hazards can be relieved at node C without incurring the added bureaucratic cost burdens (weakening of incentive intensity, added administrative costs) of unified ownership, then hybrid modes, of which franchising is an example, will be employed (provided that the contractual restrictions that accrue thereto are not treated as unlawful).

Vertical market restrictions often arise in the support of brand name capital (Klein, 1980), where the concern is that such capital will be devalued by subgoal pursuit among independent or quasi-independent distributors (often franchisees), with the result that the integrity of the system is placed at risk.\(^{10}\) Depending on the particulars of the transaction, customer and territorial restrictions, exclusive dealing, or other franchise restrictions may be imposed. Absent strategic purpose, for which pre-existing monopoly power is a requisite, the choice of instruments for imposing vertical restraints will be discerned by examining where and how the contractual hazards originate.
Price theoretic explanations for nonstandard modes of contracting include the efficiency benefits that purportedly accrue to price discrimination, the benefits of efficient risk bearing in the face of differential risk aversion, and the attenuation of free-rider hazards through the use of vertical market restrictions. The allocative efficiency benefits that accrue to price discrimination in a zero transaction cost world (which can be readily displayed in price theoretic terms) are much more problematic, however, if the costs of discovering customer preferences and of preventing arbitrage are positive. Invoking risk aversion to explain contracting practices among firms, moreover, is often second order in relation to more basic concerns with contractual hazards. Finally, unspecific free-rider claims are too often used as a shibboleth. The action, always and everywhere, resides in the details.11

4.2 the “new economy”

Is there really a new economy? Yes and no. On the one hand, there is nothing new under the sun: real time responsiveness, innovation, outsourcing, and predatory behavior are not novel issues. Each of these has been magnified, however, by the deployment of new information technologies, by an increasing appreciation for relational contracting, and by the races for the commercialization and control of information age and biotechnology developments. A change in kind seems to describe competition in many high technology sectors.

Orthodox microtheory bears on some of these issues, but often in limited ways. TCE makes limited yet productive contact in the following respects: (1) express provision for cooperative adaptation is congruent with the need for real time responsiveness; (2) innovation is examined in a systems context—in which firm size, incentives, and intertemporal transformations are featured (Williamson, 1975, pp. 196-207); (3) crafting credible commitments to support outsourcing and the bureaucratic advantages of outsourcing over internal procurement are both TCE themes; and (4) tests for predation that exonerate behavior directed at less efficient competitors (Posner, 1976, p. 193) are too static, in that they fail to make provision for contingent predation—“now it’s there, now it isn’t, depending on whether an entrant
has appeared or vanished” (Williamson, 1977, p. 339), which introduces intertemporal considerations.

To be sure, new economy issues pose strategic and knowledge creation challenges that go beyond TCE (Shapiro and Varian, 1999). Also, concepts such as “disequilibrium contracting” (Williamson, 1991) boggle the mind. That TCE is more responsive to many of the pressing needs of public policy in the new economy than is received price theory is noteworthy but scarcely grounds for complacency.

4.3 regulation/deregulation

(a) franchise bidding12

Posner’s sanguine assessment of the efficacy of franchise bidding for natural monopoly begins with the claim that to “expound the details of particular regulations and proposals…would serve only to obscure the basic issues” (Posner, 1972, p., 98). In the imperial tradition, all of the relevant action is concentrated in the ex ante bidding competition for the contract. This is consonant with Posner’s dismissive view of organization theory, to which I referred at the outset, and illustrates the pitfalls of doing public policy analysis heedless of process transformations.

Upon going beyond ex ante bidding competition to include ex post contract implementation, the attributes of the good or service to be franchised turn out to be crucial to an informed assessment. Specifically, if the good or service is to be supplied under conditions of uncertainty and if nontrivial investments in specific assets are involved, the efficacy of franchise bidding is highly problematic. The upshot is that franchise bidding for natural monopoly is not an all-purpose but is a conditional solution.13

(b) restructuring electricity supply in California

Efforts to promote efficiency by creating markets for electric power have been implemented in a number of countries with varying degrees of success. California is a recent example where the efforts to restructure have been incompletely worked through. Again, the imperial view (this is the law here) trumps the process view (what’s going on here?).
This shows up in two respects. First, “good theories” were naively expected to be implemented without making provision for the realities of the political and regulatory process. Failing to make ex ante provision for these realities, politics and regulation are conveniently made the ex post scapegoats for behaving in perverse or unanticipated ways which, in large measure, were foreseeable and should have been factored into the calculus (Williamson, 1996, Chap. 8).

Such lapses in realpolitik aside, Paul Joskow observes that too much deference was given to the (assumed) efficacy of smoothly functioning markets and insufficient attention to potential investment and contractual hazards and appropriate governance responses thereto. As Joskow puts it (2000, p. 51):

> Many policy makers and fellow travellers have been surprised by how difficult it has been to create wholesale electricity markets.... Had policy makers viewed the restructuring challenge using a TCE framework, these potential problems are more likely to have been identified and mechanisms adopted ex ante to fix them.

(c) the institutional environment

The New Institutional Economics operates at two interrelated levels: the institutional environment (or rules of the game) and the institutions of governance (or play of the game). The study of privatizing telecommunications by Brian Levy and Pablo Spiller (1994, 1996) examines the institutional environment (rules of the game) in five countries through a comparative contractual lens in which contractual hazards and credible contracting, or the lack thereof, are featured. This bottom up approach reveals that the decision to privatize and the nature of privatization vary with the condition and quality of judicial independence, the division of powers between the executive and legislative branches, the competence of the regulatory bureaucracy, and contractual safeguards. Whether and how to privatize telecommunications should therefore be made conditional on these features.
Similar considerations arise in privatizing socialist economies. The “Big Bang” approach pays little heed to differences among industries whereas those who are more concerned with cultivating institutions and the mechanisms of governance advise that a more gradual program be adopted in which the “easy cases” are privatized first. Because natural monopolies pose strains on deregulation and privatization alike (Arrow, 2000; Williamson, 2000), these are candidates to be privatized late (if at all) and then with the support of fall-back regulatory apparatus.

4.4 corporate governance/debt and equity

Price theory was long silent on the matter of corporate governance. Firms were simply assumed to maximize profits. The idea that managers might engage in subgoal pursuit that is contrary to profit maximization was inimical to the orthodox construction (although it can be and has been addressed in nearly-orthodox terms by reformulating the objective function (Baumol, 1959; Williamson, 1964)).

Transaction cost economics interprets the board of directors mainly as a security feature that arises in support of the contract for equity finance. Specifically, debt and equity are viewed not merely as alternative modes of finance, which is the law and economics construction (Easterbrook and Fischel, 1986; Posner, 1986), but also as alternative modes of governance. Thus suppose that a firm is seeking cost-effective finance for the following series of projects: general-purpose, mobile equipment; a general-purpose office building located in a population center; a general-purpose plant located in a manufacturing center; distribution facilities located somewhat more remotely; special-purpose equipment; market and product development expenses; and the like.

Suppose further that debt is a governance structure that works almost entirely out of rules. Specifically, assume that debt financing requires the debtor to observe the following: (1) stipulated interest payments will be made at regular intervals, (2) the business will continuously meet certain liquidity tests, (3) sinking funds will be set up and principal repaid at
the loan-expiration date, and (4) in the event of default, the debt-holders will exercise pre-emptive claims against the assets in question. If everything goes well, interest and principal will be paid on schedule. But debt is unforgiving if things go poorly. Failure to make scheduled payments thus results in liquidation. The various debt-holders will then realize differential recovery in the degree to which the assets in question are redeployable.

Since the value of a pre-emptive claim declines as the degree of asset specificity deepens, the terms of debt financing will be adjusted adversely. Confronted with the prospect that specialized investments will be financed on adverse terms, the firm might respond by sacrificing some of the specialized investment features in favor of greater redeployability. But then a lower cost of capital comes at an added production cost. Might it be possible to relieve the tradeoff by inventing a new governance structure to which suppliers of finance would attach added confidence? In the degree to which this is feasible, value-enhancing investments in specific assets could thereby be preserved.

Suppose arguendo, that a financial instrument called equity is invented and assume that equity has the following governance properties: (1) it bears a residual-claimant status to the firm in both earnings and asset-liquidation respects; (2) it contracts for the duration of the life of the firm; and (3) a board of directors is created and awarded to equity that (a) is elected by the prorata votes of those who hold tradeable shares, (b) has the power to replace the management, (c) decides on management compensation, (d) has access to internal performance measures on a timely basis, (e) can authorize audits in depth for special follow-up purposes, (f) is apprised of important investment and operating proposals before they are implemented, and (g) in other respects bears a decision-review and monitoring relation to the firm's management (Fama and Jensen, 1983).

The board of directors thus “evolves” as a way by which to reduce the cost of capital for projects that involve limited redeployability. Not only do the added controls to which equity has access have better assurance properties, but equity is more forgiving than debt. Efforts are
therefore made to work things out and preserve the values of a going concern when maladaptation occurs. Thus whereas the governance structure associated with debt is of a very market-like kind, that associated with equity is much more intrusive and is akin to administration. The correspondence to which I referred earlier between outside procurement/debt and vertical integration/equity therefore obtains. In effect, debt is the market form of finance and equity (the administrative form) appears as contractual hazards build up. Equity is the financial instrument of last resort.

4.5 other

Transaction cost economics maintains that any issue that arises as or can be posed as a contracting problem can be examined to advantage in transaction cost economizing terms. Accordingly, the reach of transaction cost reasoning is virtually endless. I briefly sketch two additional applications here (without bothering with price theoretic explanations).

(i) Public Bureaus. According to Douglass North, “Political markets are…prone to inefficiency” (1990, p. 365) and “high transaction cost issues gravitate to the polity” (1990, p. 372). That is worse than a paradox. That is perverse. Bad enough that political markets are inefficient. But surely the appropriate lesson is for high transaction cost issues to flee from rather than be attracted to the polity?

Maybe, but then again, maybe not. High transaction cost issues, after all, are ones that are inherently difficult to organize. As set out in Figure 1, such transactions are ones for which node A governance (in the market) is poorly suited as compared with node D governance (in the firm). If still additional contractual hazards build up, might some of these transactions be candidates for governance in the public bureau? That is precisely the argument that I advance elsewhere (Williamson, 1999). Specifically, the many disabilities of the public bureau notwithstanding—very low-powered incentives, very costly administrative procedures, very protective employment relations—there are some transactions (of which foreign affairs is an example) for which the public bureau comes off
best, judged, as it should be, comparatively. There is a place for each generic form of organization, yet each needs to be kept in its place.

(ii) Labor Organization. The organization of labor reflects many purposes, monopsony power and political purposes included. What about efficiency? Again, the action resides in the details. Those labor transactions that pose greater contractual hazards \( h > 0 \) will benefit from governance efforts to mitigate the hazards \( s > 0 \), while it will be less cost-effective to supply these same safeguards to generic labor (of a node A kind)—which is a recurrent theme. As developed elsewhere (Williamson, Wachter, and Harris, 1975; Williamson, 1985, Chap. 10), the observed organization of labor tracks an efficiency rationale.

5. Contract and Economic Organization

5.1 alternative approaches

If the contractual approach to economic organization has the reach that I ascribe to it, then the systematic application of TCE to legal education and to legal and economic research on contracting holds out considerable promise. This will entail going beyond the “sort of contract law that as flourished in American law schools: the law embodied in judicial decisions and studied by analyzing these decisions” (Rubin, 1995, p. 109). What Edward Rubin recommends instead is that the law schools (and students of contract more generally) need a “theory of contract…that addresses the contracting process itself, rather than the judicial adjudication of that process,” whereupon a “nonjudicial domain of contracting behavior” will be given prominence (1995, p. 108; emphasis added).

In principle, law and economics could have been applied to that purpose. That project, however, took a “massive wrong turn” by the argument advanced by Posner and others that “the contract law goal [of] economic efficiency…[was] achieved through common-law adjudication” (Rubin, 1995, p. 113). By drawing attention away from contracts and the contracting process toward judicial adjudication, “law and economics became just another tool for analyzing judicial decisions” (Rubin, 1995, p. 113). Rubin is nevertheless heartened that while the “law school
curriculum continues to be relatively resistant to a transactional theory of contract, ... legal scholarship has gradually begun to shift its focus as a result of the economic and sociological analysis of transactions” (1995, p. 114).

So what does a combined law, economics, and organizations approach to the study of contract, broadly conceived, entail? As I see it, the overarching move is to bring the lens of transaction cost economizing assiduously to bear. The examination of incomplete contracting in its entirety will be facilitated by supplanting the academic concept of contract as legal rules by that of private ordering and by inquiring into the mechanisms through which transaction cost economizing is accomplished. Interestingly, Ronald Gilson made many of these same arguments earlier in his examination of corporate finance transactions (1984).

5.2 the economizing perspective

The economizing perspective holds that, subject to the remediableness criterion, inefficiency invites its own demise—where inefficiency is assessed in relation to feasible alternatives (rather than a hypothetical ideal) and provision is made for implementation costs. Because joint gains will always be realized by moving from a less to a more efficient mode, provided that implementation costs do not dissipate the gains, farsighted businessmen and their lawyers will eschew inferior outcomes (such as node B in the schema). As against Machiavelli’s myopic advice to “get them before they get us,” the farsighted view of contracting is to “give and receive credible commitments” (Williamson, 1983, 1993)—by providing better information and added security features that serve to infuse confidence and realize mutual gains.

Gilson’s description of business lawyers as transaction cost engineers (1984, p. 255) is very much in this spirit. He thus urges that transactions be examined not in a one-sided way but “from the perspective of both clients” (1984, p. 245; emphasis in original), whence mutual gain is the object. He furthermore adopts a transaction cost economizing approach to private ordering (Gilson, 1984, p. 255), including express reference to credible commitments (1984, p. 281). Also, he views departures from the assumptions of the (ideal) capital asset pricing
model—namely, common time horizon, identical expectations, no transaction costs, and costless information (Gilson, 1984, p. 252)—as grist for the transaction cost economics mill: “the unreality of these [ideal]...assumptions is not cause for despair. Rather, it is the very failure of these assumptions to describe the real world that I find the potential for value creation by lawyers” (Gilson, 1984, p. 253; emphasis added). The institutions of governance arise precisely on account of these disparities (Arrow, 1963).

5.3 private ordering

(a) the concept

Marc Galanter takes exception with the usual academic/legal centralist approach to contract in which disputes purportedly “require ‘access’ to a forum external to the original social setting of the dispute [whereby] remedies will be provided as prescribed in some body of authoritative learning and dispensed by experts who operate under the auspices of the state” (1981, p. 1). The facts disclose otherwise: most disputes, including many that under current rules could be brought to a court, are resolved by avoidance, self-help, and the like (Galanter, 1981, p. 2). That is because in “many instances the participants can devise more satisfactory solutions to their disputes than can professionals constrained to apply general rules on the basis of limited knowledge of the dispute” (Galanter, 1981, p. 4). Gilson concurs: when business lawyers play the role of transaction cost engineer well, “the courts, and formal law generally, shrink dramatically in importance” (1984, p. 294).

(b) contract laws (plural)

Karl Llewellyn’s earlier dissent from the legal rules approach to contract introduces the concept of contract as framework (1931, pp. 736-737):

…the major importance of legal contract is to provide a framework for well-nigh every type of group organization and for well-nigh every type of passing or permanent relation between individuals and groups…—a framework highly adjustable, a framework which almost never accurately indicates real working
relations, but which affords a rough indication around which such relations vary, an occasional guide in cases of doubt, and a norm of ultimate appeal when the relations cease in fact to work.

This last is important, in that the prospect of ultimate appeal to the courts serves to delimit threat positions.

Related ideas have been advanced by others, including Clyde Summers who distinguishes between “black letter law” (which bears a likeness to black box economics) and a more circumstantial approach to contract. The former employs the counterfactual “illusion that contract rules can be stated without reference to surrounding circumstances and are therefore generally applicable to all contractual relations” (Summers, 1969, p. 566).

The transaction cost economics argument that each generic mode of governance is supported by a distinctive form of contract law is broadly in this circumstantial spirit. The ideal (node A) transaction in both law and economics is that of spot markets to which identity is unimportant and legal rules apply (Macneil, 1974). This legal rules approach gives way to Llewellyn’s concept of contract-as-framework as the importance of continuity builds up and incomplete long-term contracting is adopted (node C). That in turn undergoes change when transactions are taken out of the market and organized internally (node D), where the implicit law of contract now becomes that of forbearance. As previously noted, courts routinely grant standing to firms engaged in interfirm contracting should there be disputes over prices, the damages to be ascribed to delays, failures of quality, and the like, yet courts will refuse to hear disputes between one internal division and another over identical technical issues. Access to the courts being denied, the parties must resolve their differences internally (Rubin, 1995, p. 117). Accordingly, hierarchy is its own court of ultimate appeal. That firms and markets differ in their access to fiat is partly explained by these contract law differences (Williamson, 1991).

5.4 mechanisms

(a) corporate acquisition transactions
Transaction cost economics subscribes to the dictum that “explanations in the social sciences should be organized around (partial) mechanisms rather than general theories” (Elster, 1994, p. 75; emphasis in original). That is evident in the way by which TCE examines the canonical make-or-buy decision and of contracting more generally. It is also evident in Gilson’s examination of efforts by business lawyers to perfect the acquisition agreement in the face of “deviations” from the ideal assumptions of the capital asset pricing model (1984, p. 293):

- Earnout or contingent-pricing techniques respond to the failure of the homogeneous expectations assumption; controls over operation of the seller’s business during the period in which the determinants of the contingent price are measured respond to failure of the common-time-horizon assumption; and the panopoly of representations and warranties, together with provisions for indemnification and other verification techniques, respond to the failure of the costless-information.

(b) contract law doctrine

A microanalytic examination of the mechanisms that arise in conjunction with contract law doctrines would also be illuminating. Ian Macneil describes the legal system’s “less than total commitment to the keeping of promises” as follows (1974, p. 730):

Contract remedies are generally among the weakest of those the legal system can deliver. But a host of doctrines and techniques lies in the way of even those remedies: impossibility, frustration, mistake, manipulative interpretation, jury discretion, consideration, illegality, duress, undue influence, unconscionability, capacity, forfeiture and penalty rules, doctrines of substantial performance, severability, bankruptcy laws, statutes of fraud, to name a few; almost any contract doctrine can and does serve to make the commitment of the legal system to promise keeping less than complete.
The refusal by the courts to enforce stipulated damages clauses is especially puzzling. Since the parties to a contract can be presumed to know best what contractual terms serve their interests, why should the courts refuse to enforce stipulated damages in the event of breach?

One possibility is that contract is a devious thing. Thus although such a clause may frequently be the efficient way to settle a breach, it could also serve strategic purposes, of which induced breach is one.

The issue of contrived cancellation has been addressed by Kenneth Clarkson, Roger Miller, and Timothy Muris in their discussion of refusal of the courts to enforce stipulated damage clauses where breach has been deliberately induced (1978, pp. 366-372). Induced breach could arise where a party intentionally withholds relevant information, yet complies with the letter of the contract. Or it might involve perfunctory fulfillment of obligations where more resourceful cooperation is needed (1978, pp. 371-372). In either case, induced breach is costly to detect and/or prove (1978, p. 371). Transaction cost considerations are plainly operative.

5.5 ramifications for legal education

Gilson advises that my observation that the legal centralism approach to contract relieves “lawyers and economists...of the need to examine the variety of ways by which individual parties to exchange ‘contract out of or away from’ the governance structures of the state by devising private orderings” (Williamson, 1983, p. 520) is too sweeping. It should be restricted to academic lawyers and economists (1984, p. 295). That is because “business lawyers have done an awfully good job at something the law schools did not and, for the most part, still do not teach: helping people arrange their relationships in the absence of governmental intervention: facilitating private ordering” (Gilson, 1984, p. 303; emphasis in original). But then “why have law schools done so bad a job training business lawyers?” (Gilson, 1984, p. 303). Gilson’s answer is that “There has been no theory...that dealt with private ordering” (1984, p. 304) prior to the appearance of “two areas in economics—finance and transaction cost economics” (1984, p. 305).
Sixteen years later we find that the teaching of contract law has changed very little. What explains this continuing neglect?

One explanation is that mainline law and economics has remained comfortably ascendant. The relation between law and economics thus continues to be one in which textbook economic orthodoxy is the fount. The predilection to work out of a theory of the firm-as-production function setup is thus reaffirmed and the subject of organization remains disjunct. Reservations about the efficiency of common law adjudication notwithstanding, contract law teaching stays predominantly focused on legal rules and adjudication.

A second explanation is that the world of private ordering is impossibly complex. Good lawyers being quick studies, better that they learn about private ordering on-the-job rather than in the classroom.

The first of these arguments is a lame excuse for complacency, while the second overlooks the possibility that the economics of organization involves variations on a few key themes. In that event, attention can be focused on canonical cases—of which credible interfirm contracting is one and vertical integration is another. The buzzing, blooming confusion of private ordering is thereby reduced to more manageable proportions. Since the classroom is the place to lay out the intuition, merits, and mechanisms of credible (node C) contracting and to examine the comparative strengths and weaknesses of the firm-as-governance structure (node D), to relegate the study of private ordering to on-the-job training is anachronistic.

Even, moreover, if the basic law school curriculum is unmoved by these arguments, it is noteworthy that a number of leading law schools have begun to offer an elective course on complex “deals,” many of them modelled after the course offered by Gilson and Victor Goldberg at Columbia Law School on “Deals: The Economic Structure of Transactions and Contracting.” If the demand for transaction cost engineers cannot be met by the law schools, the business schools could end up eating that lunch (Rubin, 1995, p. 114).

6. Conclusions
There is growing agreement that “the objectives of firms, the reason for their existence and the manner of their decision taking…will require modes of analysis quite different from those which have dominated in this century” (Hahn, 1991, p. 49). Not only does transaction cost economics hold that the way to think about contract and organization is to bring the purposive and farsighted lens of economizing to bear, but the existence and governance of firms are both the key TCE issues.

As developed herein, organization theory has massive ramifications for the TCE theory of the firm. Salient contributions from organization theory include the description of human actors in more veridical terms, the importance of intertemporal process transformations, choice of the unit of analysis, and the description of alternative modes of governance as syndromes of complementary attributes. The resulting theory of the firm differs greatly from the neoclassical (Kreps, 1990, p. 96). Because “Any standard theory, not just neoclassical, starts from the existence of firms” (Arrow, 1999, p. vii), that is very basic.

To be sure, the proximate lessons (as advanced by organization theorists) and the ultimate lessons (as viewed from an economizing perspective) often differ—and that is consequential. But the more basic point is this: someone needed to step up and offer trenchant critiques and identify relevant phenomena. Organization theorists were prepared to do that when others were complacent or held back.

The theory of the firm-as-governance structure that is sketched herein is an ongoing rather than finished construction. Its evolving status notwithstanding, it has already served to deepen our understanding of many complex contractual and organizational phenomena and operates as a check against overuses and misuses of orthodoxy. It is in that spirit that I recommend that mainstream law and economics stands to benefit by incorporating the lessons and some of the methods of law, economics, and organization—both as these bear on public policy and in relation to the law school curriculum.
Footnotes

“The author is Edgar F. Kaiser Professor of Business Administration, Professor of Economics, and Professor of Law at the University of California, Berkeley. This paper was originally prepared for the ceremony inaugurating the Program in Law and Economics at the University of Chile in August 2000. That Program is a joint effort between the Faculty of Economics and Business and the Faculty of Law to “promote and develop interdisciplinary theoretical and applied research in the area.” It was subsequently presented at the opening session of the 4th annual conference of the International Society for New Institutional Economics in Tuebingen, Germany, at the Law and Economics Seminar at George Mason University, and the Law, Economics, and Organization Seminar at the University of Southern California. Useful questions and comments at all four sessions are gratefully acknowledged.

1. Posner’s famous text is titled “economic analysis of law,” and he takes the (one-way) position that the “economic analysis of law…is what ‘law and economics’ means” (Posner, 1993, p. 83).


3. The businessman Rudolf Spreckels put it in an entrepreneurial way: “Whenever I see something badly done or not done at all, I see an opportunity to make a fortune.” Such entrepreneurs will both see opportunities and seize the initiative. Although I would expect that some of the concepts and apparatus from the modelling of governance will carry over, modelling entrepreneurship has so far proven to be very difficult.

4. Organization theory is a huge and varied subject. I draw upon it selectively, relying principally on the lessons of Carnegie (Simon, 1957a, 1957b; March and Simon, 1958; Cyert and March, 1963; Simon, 1997).
5. Although the nonverifiability of contracts is widely conceded, inefficiency in the ex post implementation stage of contracting can be annihilated by assuming that the principals costlessly bargain to the efficient result in all state realizations. In that event, all of the analytical action is concentrated in the ex ante stage of contracting (Hart, 1995). As I have observed elsewhere, the assumption of costless bargaining is a preposterous simplification (Williamson, 2000).

6. Neoclassical theory does make provision for externalities, which can be thought of as unanticipated consequences. But these are rarely of the organizational kinds to which the organization theory literature directs our attention.

7. It is uncontested that Coase (1937) is a pathbreaking article. It is unrealistic, however, to expect that such an article, which opened up a vast new area for research, should “do it all.” As Coase has subsequently stated, the reason why the article remained in a nonuse status until the 1970s was because the key ideas awaited operationalization (Coase, 1992, p. 719).

8. Surveys of empirical transaction cost economics by Howard Shelanski and Peter Klein (1995), Bruce Lyons (1996), Keith Crocker and Scott Masten (1996) and Aric Rindfleisch and Jan Heide (1997) are pertinent, as is a more recent review of the relative contributions of agency theory and TCE to the empirical literature on contracting. Whereas TCE has advanced numerous refutable implications, agency theory is notable for “its failure to generate testable hypotheses…[or even] to account for…the more basic features of real world contracts” (Masten and Saussier, 2000, p. 110).

A survey of empirical TCE currently in progress (Boerner and Macher, 2000) estimates that the cumulative number of empirical TCE studies has grown from 200 in 1994 to over 600 in the year 2000. Econometric refinements, moreover, are being introduced as this work has progressed (Masten and Saussier, 2000). To be sure, transaction cost economics, like everything else, will benefit from more and better
empirical work. I have no hesitation, however, in declaring that transaction cost economics is an empirical success story. Judged comparatively, “this empirical work is in much better shape than much of the empirical work in industrial organization generally” (Joskow, 1991, p. 81).

9. Somewhat more formal reduced form models that are broadly in the spirit of the Simple Contractual Schema include Michael Riordan and Williamson (1985) and Williamson (1991).

10. See the quotation to Coase (1972) in the text of Section 1, supra. Differential risk aversion supplanted monopoly as the favorite price theoretic explanation for nonstandard practices in the 1970s.

11. Although Posner contends that “the proper lens for viewing antitrust problems is price theory” (1979, p. 932), Alan Meese observes that “Despite references by Chicagoans to ‘price theory,’ Chicago’s approach to vertical restraints has never rested upon…price theory. Instead, the Chicago approach to vertical restraints is an application of [NIE/TCE reasoning]” (1997, p. 203). Also see Joskow (1991, pp. 567-57).

12. This subsection is elaborated in Williamson (1996, pp. 84-85).

13. Examples where franchise bidding for goods and services supplied under decreasing cost conditions can possibly supplant extant regulation or public ownership with expected net gains include local service airlines and, possibly, postal delivery. The winning bidder for each base plant (terminals, post office, warehouses, and so on) can be owned by the government, and other assets (planes, trucks, and the like) will have an active second-hand market. It is not, therefore, that franchise bidding is totally lacking in merit. On the contrary, it is a very imaginative proposal. Transaction cost economics maintains, however, that all contracting schemes—of which franchise bidding for natural monopoly is one—need to be examined microanalytically and assessed in a comparative institutional manner.
14. This subsection is based on Williamson (1996, pp. 184-185).

15. Farsighted contracting is more plausible in intermediate product market contracts than in final goods markets. Still, farsighted firms that are selling to consumers who lack the relevant expertise and foresight nevertheless can and do take steps to alleviate the hazards—through branding, warranties, guarantees, and the like. I do not mean to suggest, however, that here is never an occasion to craft additional relief (possibly with the aid of public policy) against residual hazards.

16. Full formalization is the ultimate objective. The Grossman-Hart-Moore model (Hart, 1995) qualifies as a fully formal model but is lacking in plausibility (Kreps, 1999). The recent treatment of procurement by Patrick Bajari and Steven Tadelis (2000), which focuses on the incentive and ex post adaptation differences between fixed price and cost plus contracting, is much closer in spirit to TCE.

17. One of the comments that I have received on this paper is that the basic message has not only been heard, but that it has registered and taken effect. That is gratifying, yet other readers remark that much of this is unfamiliar terrain and needs to be more fully spelled out.

I come out somewhere in between. Thus although many of the firm-as-governance structure ideas have taken hold, private ordering remains underdeveloped and organization theory is scanted by mainline law and economics—witness the leading textbooks (Cooter and Ulen, 2000; Polinsky, 1989; Posner, 1998). Public policy inroads notwithstanding, the basic contract law course remains immune to the arguments in this paper.
References


